

TR

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 4 JAN 2002		3. REPORT TYPE AND DATES COVERED Final Report
4. TITLE AND SUBTITLE Unit climate, leadership, and performance: An aggregate-level investigation			5. FUNDING NUMBERS	
6. AUTHOR(S) CPT Jeffrey L. Thomas, MAJ Paul D. Bliese, LTC Craig Bullis				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Walter Reed Army Institute of Research Washington, DC 20307 Center for Army Leadership Fort Leavenworth, KS			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) US Army Medical Research and Materiel Command Fort Detrick, MD			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release, distribution unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) This study examined the link between stressors, strains and performance among 31 company sized units. In addition, the 360 degree assessment tool of leadership was examined. Results revealed several key linkages between stressors, strains and unit performance. The evaluation of the 360 degree assessment tool, however, revealed that the instrument had poor discriminant validity.				
<div style="font-size: 2em; font-weight: bold;">20020520 208</div>				
14. SUBJECT TERMS Stress, Leadership, Performance, 360 Degree Evaluation			15. NUMBER OF PAGES 46	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT	



***Unit Climate, Leadership, and Performance:
An Aggregate-Level Investigation***

CPT Jeffrey L. Thomas, Ph.D.

MAJ Paul D. Bliese, Ph.D.

Department of Operational Stress Research
Walter Reed Army Institute of Research
Washington, DC

LTC Craig Bullis, Ph.D.

Center for Army Leadership
Fort Leavenworth, KS

4 JAN 2002

NOTE: Comments should be directed to CPT Jeffrey L. Thomas at the United States Medical Research Unit-Europe, Nachrichten Kaserne, Karlsruher Str. 144, 69126 Heidelberg, Germany, telephonically: +49-6221-172007, electronically: jeffrey.thomas@hbg.amedd.army.mil

The views expressed in this report are those of the authors and do not necessarily represent the official policy or position of the Department of Defense, Department of the Army, Army Medical Department, Army Medical Research and Materiel Command, Walter Reed Army Medical Center, Walter Reed Army Institute of Research, Center for Army Leadership, or the U.S. Government.

ACKNOWLEDGEMENTS

The authors wish to thank the Officers and soldiers of the participating Field Artillery Brigade and Armored Cavalry Regiment, and particularly the Commanders and S-3s for each of these units. Your help and support were greatly appreciated and we hope to be able to work with your units in the future!

We would also like to acknowledge LTC Rick Fagan and SFC Bruce from the Center for Army Leadership for their hard work, guidance, and expertise in coordinating and completing the project.

Lastly, we would like to thank CPT Darren Ritzer, SSG John Eline, CPL Michele Gold, and Ms. Wanda Cook from the Department of Operational Stress Research at the Walter Reed Army Institute of Research for their professional and technical expertise in all phases of this project.

1	EXECUTIVE SUMMARY	5
2	INTRODUCTION	7
2.1	Background.....	7
2.2	Study Objectives	8
2.3	Working Research Model	9
2.4	360-degree Assessment.....	11
2.4.1	What is it?	11
2.4.2	Who does it and why?.....	11
2.4.3	Leader Performance: It depends on who you ask	11
2.4.4	Congruence among rating sources	12
2.5	Stressors.....	14
2.5.1	Role Stressors	14
2.5.2	Interpersonal Conflict	15
2.5.3	Situational Constraints	15
2.6	Strains (Antecedents of Performance)	15
2.7	Measuring Performance	16
2.7.1	Performance, Behavior, Effectiveness and Productivity.....	17
2.7.2	Enlarging the Criterion Space: Dimensions That Matter	17
2.8	Leader Moderators.....	19
3	METHOD	20
3.1	Sample	20
3.2	Procedure	20
3.3	2.3 Scales and Data used to Assess Model Variables	21
3.3.1	360-degree Leadership Ratings--CAL's Azimuth	21
3.3.2	Individual Difference Moderators.....	22
3.3.3	Stressors	22
3.3.4	Strains	22
3.3.5	Archival Performance and Scale-Based Performance Measures	22
3.4	Plan for Data Analyses.....	22
4	RESULTS.....	23
4.1	Means and Standard Deviations for all Study Variables.....	23
4.2	Correlational Analyses: Stressors, Strains, and Performance	23
4.2.1	Stressors and Strains	23
4.2.2	Stressors and Performance	24
4.2.3	Strains and Performance	26
4.3	Correlational Analyses: Rater Congruence, Ratings and Performance.....	26
4.4	Exploratory Factor Analysis: The Factor Structure of the Azimuth	26
5	DISCUSSION.....	27
5.1	Summary of findings.....	27
5.1.1	Unit Climate and Unit Performance.....	28
5.1.2	360 leadership ratings and unit performance	29
5.1.3	Psychometric properties of the Azimuth.....	30
5.2	Future Work.....	30
6	REFERENCES	32
7	TABLES	35
7.1	Table 1: Descriptive Statistics for Aggregate Level Study Stressors (n = 31 Companies).	35
7.2	Table 2: Descriptive Statistics for Aggregate Level Strains/Performance Antecedents (n = 31 Companies). 36	
7.3	Table 3: Descriptive Statistics for Aggregate Level Performance Dimensions (n = 31).	37
7.4	Table 4: Means and Standard Deviations (Standardized) for all 360-degree Azimuth Dimension x Rating Source 38	
7.5	Table 5: Correlations between Aggregate Level Stressors (1-6) and Strains (7-12) (n = 31).	39
7.6	Table 6: Correlations between Aggregate Level Stressors (1-6) and Performance (7-13) (n = 31).....	40
7.7	Table 7: Correlations between Aggregate Level Strains and Performance (n = 31).....	41

7.8	Table 8: Correlations between Global 360-degree Assessment by Rating Source and Performance (n=31)	42
8	APPENDIX A: Leadership Dimensions	43
9	APPENDIX B: Items from CAL Azimuth 360-Degree Assessment Tool	45

1 EXECUTIVE SUMMARY

Few would argue that it is critical for Army units to understand how stress affects the performance and health of its soldiers. In fact, this has become a particularly salient issue recently as Army researchers have begun to examine organizational stressors unique to the military such as operational tempo (OPTEMPO), personnel tempo (PERSTEMPO), and deployment stressors. Given the focus on these stressors, it becomes important to understand not only the impact of stressors on soldier health and soldier strains such as job satisfaction and commitment, but also the impact of stressors on objective measures of soldier/unit performance. The simple assertion that soldier/unit health and organizational strains such as commitment are related to performance outcomes is certainly implied by lay people and researchers alike. However, to our knowledge this link has not been explicitly examined before.

Furthermore, leadership has always been, and will continue to be a top priority for the Army in terms of the training and development of its soldiers. However, just as we noted the lack of research addressing the stress-performance link, there has been very little research examining the impact of leadership on important soldier/unit outcomes. Thus, we feel that the impact of leadership on individual and unit performance has been under-represented in the literature. The present study examined these links in some detail.

In order to address these concerns, the Department of Operational Stress Research at the Walter Reed Army Institute of Research (WRAIR) collaborated with the Center for Army Leadership (CAL) on a large research project. The research project was designed to examine the links between *unit climate measures of stress and strains* (health and organizational), *performance* (archival and survey-based), and *360-degree evaluations* (supervisory, peer, self, subordinate) of Army leaders on doctrinally based dimensions relevant to the Army. These data were collected and aggregated for 31 company-sized units and 2,400 soldiers from Fort Sill, OK and Fort Irwin, CA.

The study's key contribution is that it represents the first use of broad-based survey data assessing unit climate, unit performance (archival and survey-based), and 360-degree leadership assessments in a single study. The study assesses these relationships using a *multi-method* (self-report survey data, objective data, multi-rater survey data), *multi-trait* (leadership dimensions, stressor, strains, and unit measures of performance) approach.

The study's key objective was twofold--examine the links between, 1) unit climate (stressors and strains) and unit performance, and 2) 360-degree assessments and unit performance. The secondary research objective was to assess the psychometric properties of the 360-degree leadership tool by testing the instrument's reliability, validity, and factor structure.

In general, limited support was found for the link between unit climate and unit archival performance with the exception of work-family conflict and APFT scores, and UCMJ Offenses and Financial Problems. Grouped sample sizes in the present study were quite small ($N=31$), thus the stability of these findings and of the non-significant links observed between unit climate and unit archival performance must be viewed with caution.

Conversely, the linkage between climate and unit performance was demonstrated using survey-based measures of performance. A key survey performance outcome was commitment (affective and continuance). Commitment was the outcome most strongly related to unit climate stressors and unit strains. Combat readiness, a survey performance measure of great interest to Commanders, was significantly related to horizontal cohesion, vertical cohesion for Officers and NCOs.

The linkage between 360-degree leader assessment and unit performance was also validated. Of the multi-rater assessments, peer ratings had the most significant correlations with performance (e.g., M16 scores, APFT scores, and Organizational Citizenship Behavior). One final key finding regarding the multi-rater leader assessments involved appraisals of combat readiness; both subordinate and supervisory leader ratings were significantly and positively related to combat readiness. The highest rater congruence was between the subordinate and supervisory rating groups.

Lastly, we sought to study the psychometric properties of the 360-degree assessment tool developed by the Center for Army Leadership, the Azimuth. Our analyses revealed that there was little discriminant validity among the scales making up the Azimuth. That is, raters tended to rate leaders on one "global factor" instead of the dimensions specified in the Azimuth. If a leader was effective, he or she tended to be rated highly on all the factors, therefore, there was little differentiation between the sub-scales of the Azimuth.

In sum, our objectives were met. Valuable information about the link between stressors/strains and unit performance (archival and survey) was identified. In addition, the 360-assessment tool yielded results that were predictive of unit performance. However, it should be noted that future work should address the factor structure and underlying dimensionality of the Azimuth. Given the relationships observed in the present study, we suggest that future researchers focus on the psychometric properties. Its utility as a developmental tool and as a research tool is promising.

2 INTRODUCTION

2.1 Background

Effective leadership is, and always will be, critical to the success of the Army. Consequently, the Army continues to be proactive in its efforts to provide leadership development tools. One such developmental tool is a "360-degree" feedback instrument. 360-degree feedback is conducted by having a target leader assess his or her strengths and weaknesses along a number of dimensions relevant to leadership performance. The leader's self-evaluation is conducted in conjunction with peer, subordinate, and supervisor evaluations, which rate the target leader's strengths and weaknesses along these same dimensions. Thus, the term "360-degree" refers to the multiple sources for leadership evaluations, which represent feedback from all levels in the organization. Specifically, feedback consists of (a) assessments of the leader's strengths and weaknesses from these multiple sources, and (b) comparisons of the congruence between self and other ratings (e.g., peer, subordinate, supervisor).

Currently, the Center for Army Leadership (CAL) at Fort Leavenworth is the Army's proponent for designing and field-testing a 360-degree feedback instrument aptly called the "Azimuth". To date, CAL has successfully conducted several 360 evaluations of the Azimuth with Battalion and Brigade-sized units. As the next step in the development of the Azimuth, CAL was interested in examining linkages between Azimuth evaluations and unit climate and performance. Specifically, researchers from CAL were interested in examining whether leaders who receive high ratings on the Azimuth command units with a strong positive climate that are also high performing. In this way, the 360-degree assessment process (Azimuth) could be validated against unit climate and unit objective measures of the rated leaders.

Over the years, the Department of Operational Stress Research (DOSR) at Walter Reed Army Institute of Research (WRAIR) has developed expertise in quickly and reliably assessing several important dimensions of unit performance. The bulk of this research has examined performance by assessing unit climate in terms of stressors and modeling the negative effects of stressors in terms of unit and soldier strains. Data gathered by the DOSR has historically been analyzed for two purposes: 1) to consult the Army commanders on issues affecting soldier and unit health and performance, and 2) to develop operational stress models across military settings (e.g., deployments, training, garrison). By developing predictive models through the use of unit climate assessment, the DOSR has been able to measure the impact of operational stress in order to help support, protect, and sustain soldier health and performance at the individual and unit level.

Although DOSR researchers infer that poor unit climate (e.g., high stressors and high strains) has deleterious effects on individual and unit performance, this link is not well established by researchers in either military or academic circles. The present study represented an attempt to link unit objective measures to both 360-degree leadership and unit climate assessment.

2.2 Study Objectives

This collaborative research project between the Center for Army Leadership and Walter Reed Army Institute of Research had three key objectives.

- To assess the impact of unit climate (unit stressors and strains) on dimensions of unit performance. It is hypothesized that high stressor and strains will generally manifest themselves in lower performance. As state above, this link has received little attention in previous research; therefore, this study represents an important step in assessing the link between unit climate and “hard” performance dimensions. Because performance is believed to be multi-dimensional in nature, it is our assertion that specific stressors and strains will impact performance differentially. Note that to test such models, both unit climate and objective measures of performance must be based at the same aggregate level (i.e., the company level).
- To determine the relationship between multi-rater leadership assessment on dimensions of performance using objective unit measures and survey-based measures. It is hypothesized that different sources of leadership ratings (e.g., subordinate, peer) will show differential predictive relationships with the unit objective measures assessing dimensions of unit performance. This stems from the long-held assertion that performance is a multi-dimensional construct (see Campbell, 1999) and that different rater groups will perceive certain aspects of performance as important while other rater groups will not (see Bozeman, 1997; Harris & Schaubroeck, 1988; Mount, Judge, Scullen, Systma, & Hezlett, 1998). It is important to note here that the unit climate data has been aggregated to the company level. Thus, each Company Commander ‘s individual 360-degree data has been linked to his/her units’ climate assessments.
- To validate the Center for Army Leadership’s 360-degree Azimuth tool. CAL is the Army’s lead agency in developing and implementing multi-rater assessment via 360-degree ratings. Because the Azimuth is still being developed, CAL can validate its predictive and construct validity against unit climate and unit objective measures. The data gathered by WRAIR can aid to the Azimuth’s continual development as a leadership tool. This study will be of great benefit to CAL by allowing for the study of the psychometric properties of the instrument. Once the psychometric properties are assessed the Azimuth can become more refined, and thus be a valid and reliable means for the Army’s important mission of developing leaders. Moreover, knowledge gained from the study on the psychometric properties of CALs Azimuth will aid researchers at WRAIR in the development of a Leadership Climate Inventory (LCI). Valid and reliable factors that emerge from the Azimuth will be added to our pre-existing leadership measures to form the LCI. In this way we can augment the Vertical Cohesion scale of leadership currently used at WRAIR by identifying additional discriminators of leadership behavior as assessed from multiple sources.

Having laid out the study objectives, it is now necessary to frame our discussion of the key components of the study by making use of a working research model. We will present the basic research model used for the present study and describe the overall schematic. We will then turn to a literature review on the following key components of the model: 1) 360-degree assessment,

2) stressors, 3) strains--antecedents of performance, and 4) performance dimensions. Next, we will describe our methodology, data collection, and plan for data analyses. Lastly, we will present our results followed by a discussion of the implications of our findings.

2.3 Working Research Model

Figure 1.1 summarizes the working model for the project. Since the research project makes use of individual-level and aggregate-level data, and for the sake of clarity, we need to point out what particular data were measured at the individual level and what data were measured at the company level.

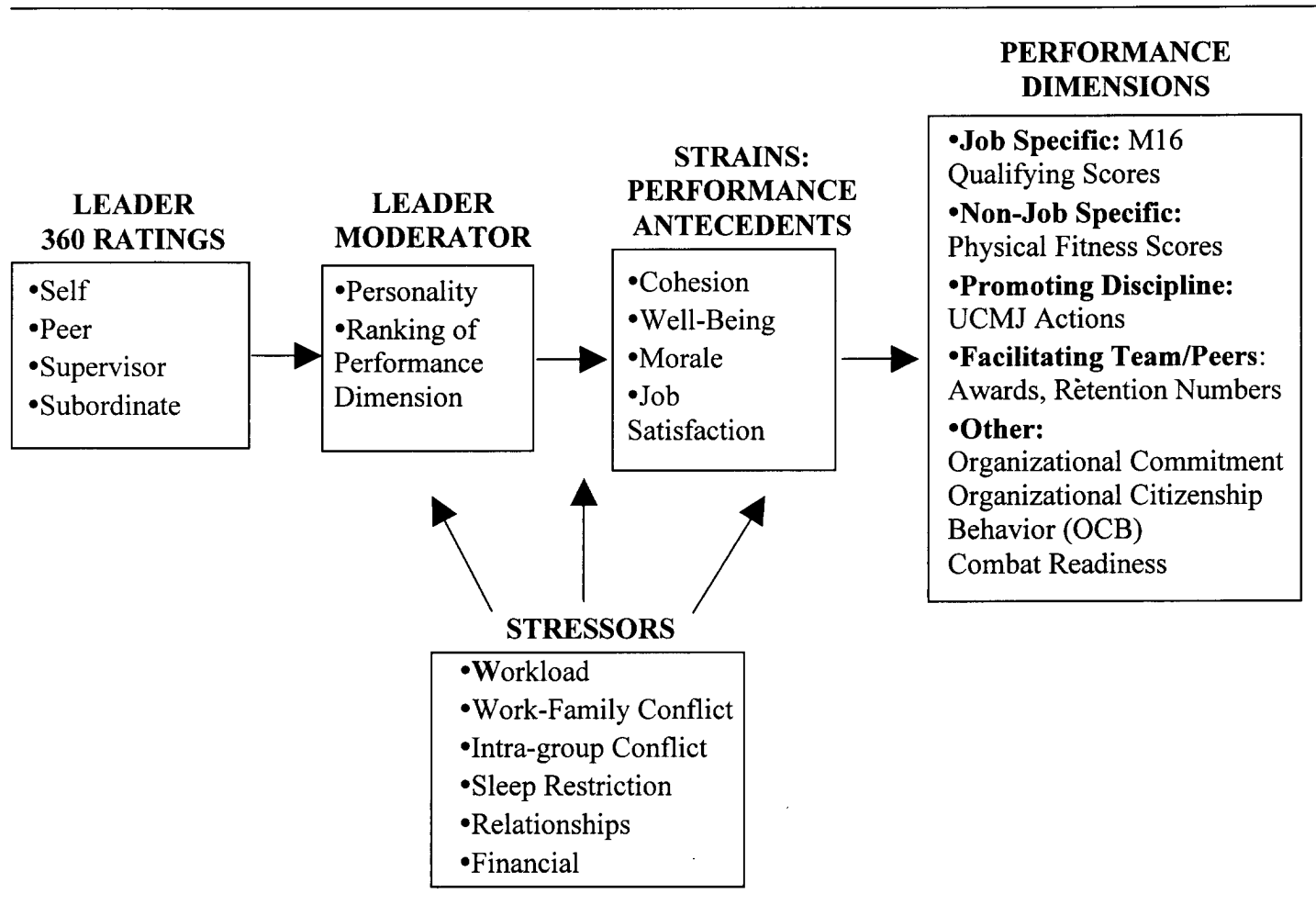
The 360-degree ratings and the leader moderator data were assessed at the individual level of analysis. However, it is important to note that the 360-degree rater groups of peers, subordinates, and in some cases supervisors had more than one rater per group. For example, a target leader may have had six peer raters. Thus, the target leader would have had six peer 360-degree assessments. Data such as these were aggregated to form one rating per source (e.g., peer, subordinate). There were four sources of 360-degree ratings: self, peer, subordinate, and supervisory.

The dimensions of Army leadership assessed by CALs 360-degree tool--Azimuth are as follows: 1) communication, 2) decision-making, 3) motivating, 4) developing, 5) building, 6) learning, 7) planning and organizing, 8) executing, 9) assessing, 10) respect, 11) selfless service, 12) integrity, 13) physical attributes, 14) overall leadership, 15) tactical skills, and 16) technical skills. CAL assessed these dimensions since they tapped themes of the familiar Be-Know-Do Model of leadership put forth in FM 22-100 (see Appendix A) along with markers of performance generally assessed in the Officer Evaluation Reporting System (OERS--see AR 623-105). Furthermore, by being a part of the FM 22-100, Army Leadership as well as the OER system, these dimensions have been institutionalized as aspects of leadership performance that are at the very core of successful leadership for the Officers corps. Items used to tap each dimension are presented in Appendix B.

The leader moderator data consisted of the individual leader's responses, i.e., the 30 Company Commanders' responses, to the personality measures of Type A, Self-efficacy, Conscientiousness, Neuroticism, and the leader's subjective rank-order of performance dimensions proposed by Campbell, McCloy, Oppler, and Sager (1993). The personality variables listed above, each, have demonstrated relationships with performance appraisal in past research (e.g., Barrick & Mount, 1990; Thomas, 1999).

Unit stressors, strains--antecedents of performance, and performance dimensions were assessed at the company level. Specific stressors surveyed include: intra-group conflict, work-family conflict, work overload, work hours, sleep, marital/relationship problems, financial problems, and work unpredictability. These stressors have demonstrated strong relationships with unit strains in past research conducted by the WRAIR as well as the academic literature.

Figure 1.1: CAL/WRAIR Project Model



Antecedents of performance assessed were: vertical cohesion (officer and non-commissioned), horizontal cohesion, general well-being, morale, and job satisfaction. Because this study represents our first use of “hard” performance measures, we thought it useful to characterize variables we typically examine as unit strains, as antecedents of performance. Since the testing of the link between strains and performance is a key objective here, this was felt necessary to fit with the literature on measuring performance. That is, we suggest that what is referred to as a *strain* in stress models can be seen as an *antecedent of performance* in performance models (e.g., Muchinsky, 1990). We will discuss this in more detail below.

Dimensions of performance were assessed by examining archival performance records and by survey. The archival data tapped into dimensions of performance originally proposed by Campbell et al. (1993). Specifically, these included weapon qualification scores, physical fitness scores, UCMJ actions (violations against the uniform code of military justice), number of awards given, and retention numbers for first-term and mid-term soldiers. Survey-based performance measures consisted of organizational citizenship behavior, organizational commitment (affective

and continuance), and soldiers' perceptions of unit combat readiness. Greater detail for each of the measures is listed below in the Methods section.

2.4 360-degree Assessment

2.4.1 What is it?

Because 360-degree feedback is a new process to the Army, we will review the extant literature on its use in some detail. 360-degree feedback is a performance appraisal process in which employee evaluations are collected from multiple sources. The term "360-degree" is associated with multi-rater assessments that get evaluations from supervisors, subordinates, peers, and the self; therefore, employees are evaluated in a concentric fashion within the organization's hierarchy. 360-degree evaluations can be seen as an extension of the more traditional performance appraisal using supervisory evaluations only.

2.4.2 Who does it and why?

The use of multi-rater assessments has gained considerable acceptance over the past two decades. In fact, London and Smither (1995) recently conducted a survey with leading consultant firms and found that multi-source feedback is widely used. One of their survey respondents stated that its use "was nearly universal...every Fortune 500 firm is either doing it or thinking about it." It has been used for both developmental and administrative purposes within organizations. However, it is important to point out that its implementation has been viewed more positively by employees when used as a developmental tool as opposed to an administrative one (see Bettenhausen & Fedor, 1997).

While the use of multi-source feedback by practitioners and organizations has flourished, this important development has not received an equal amount of attention by researchers in industrial/organizational psychology. That is, it would seem that practice is well ahead of empirical research (London & Smither, 1995). However, researchers are beginning to pay more attention to multi-rater assessment and have demonstrated numerous advantages associated with the use of multiple raters in the evaluation of a single employee including: enhanced ability to observe and measure various job facets (Borman, 1974), greater reliability, acceptance and perceived fairness (Latham & Wexley, 1982; Bettenhausen & Fedor, 1997), increased accuracy (Zammuto, London, & Rowland, 1982), improved legal defensibility via Title VII of the Civil Rights Act of 1964, and its benefits over single-source data (Bernardin & Beatty, 1984; Cascio & Bernardin, 1981).

2.4.3 Leader Performance: It depends on who you ask

The use of multi-rater assessment has become popular because organizations are beginning to recognize that performance is not unidimensional in nature, but rather multidimensional. In essence, multi-rater assessments broaden the criterion domain by taking into account that raters at different levels in the organization will observe different facets of a target leader's job

performance. Thus, a target leader's performance is likely perceived differently depending on the source of the ratings. Feedback from multiple sources provides a target leader with information not otherwise available using supervisory evaluations alone.

One can readily see the implications of the multi-rater approach; varying aspects of leader performance are important to subordinates (whom see the target leader as a supervisor), supervisors (whom see the target leader as a subordinate) and peers (whom see the target leader as a colleague). This implicit understanding of the multidimensional nature of leader performance has led to the increased use of multi-rater assessment by organizations for individual leader/manager development and feedback. Practitioners have realized that this type of feedback is vital for target leaders to receive. London and Smither (1995) point to a conceptual reason for the value a target leader may gain from this type of feedback. They note that the work environment is, in essence, a socially constructed world in which employees at all levels interact. Naturally, others' judgments about a target leader are part of the social reality, whether biased or not, and it would serve one well to attend to as such.

The conceptualization of the value a target leader may gain from multi-source feedback leads to the question of how much do raters agree about a target leader. That is, since we know that employees will see varying aspects of a leader's performance as critical or desirable depending on the rating source, how much congruence is there between and within rating sources (e.g., subordinates, supervisors, peers).

2.4.4 Congruence among rating sources

*Oh wad some power the giftie gie us
To see oursel's as other see us!
It would frae mony a blunder free us,
And foolish notion--Robert Burns*

The above, quote by Burns (as cited in the Army's Officers Guide, 45th edition), illustrates that "seeing ourselves as others see us" would give us keen insight into our own actions. Multi-rater assessment gets right at the heart of Burn's prose by allowing a target leader to see their own actions at work through the eyes of those around them. (360-degree.) As noted above, there seems to be no doubt as to the advantages of using a multi-rater assessment approach over the single supervisor-rater assessment. However, Burns playful suggestion is not so easy to pull off. Researchers have turned their attention to a rather unique dilemma that the development and wide use of multi-rater assessments has created. This dilemma has to do with the congruence, or lack of congruence, between different rating sources. That is, there is debate among researchers about the extent to which we should expect high inter-rater agreement between the rating groups (e.g., peers, subordinates, supervisors, and self). Is it possible to see ourselves as others see us?

In an attempt to answer this question, Harris and Schaubroeck (1988) reviewed the extant literature examining self-peer, peer-supervisor, and self-supervisor ratings. Subsequently they performed a meta-analysis to synthesize their results. They found that the highest congruence existed between the peer-supervisor dyad ($p = .62$), while there was relatively modest

congruence between dyad involving self-ratings (self-supervisor, $\rho = .35$; self-peer, $\rho = .36$). So, it can be gleaned that there some degree of agreement between others' ratings (ratings between peers and supervisors), however, there is much less agreement in self-other rating dyads.

Harris and Schaubroeck's (1988) explanation points to two versions of egocentric bias on the part of self-raters. First, self-ratings may be more lenient because a self-rater is inclined to inflate his/her rating in order to enhance the evaluation (Holzbach, 1978; Steel & Ovalle, 1984). In turn, this leads to restriction of range and attenuation of the correlation between self-other rating pairs. Conversely others' ratings would not suffer from range restriction because there would not likely be a leniency effect when rating others.

Secondly, classic social psychology tells us that all of us are prone to commit the fundamental attribution error repeatedly (Jones & Nisbett, 1972; Ross, 1977). The fundamental attribution error is the tendency for actors to attribute positive events to their own characteristics/disposition and negative events to the situation. The flip side of this, of course, is that observers are more likely to attribute positive events for the person to the situation, and attribute negative events for the person to characteristics/disposition of that person (this conceptualization is quite similar to Fiske and Taylor's (1991) notion of actor-observer differences). Both leniency effects and the fundamental attribution error are plausible explanations for why self-ratings perform so poorly.

In addition to the finding that there was less agreement between self-other rating dyads, Harris and Schaubroeck's review also found support for the leniency bias/egocentric bias. When comparing mean average rating for each source (peer, supervisor, self) at the global level, they found that global self-ratings were over a half a standard deviation higher than supervisor ratings and one quarter of a standard deviation higher than peer ratings.

Furnham and Stringfield (1994) conducted a study further examining congruence between multi-rater sources. They found that congruence was much higher between supervisors and subordinates than self with either source. Peers were not assessed, but the results are consistent with the findings of Harris and Schaubroeck (1988). In a subsequent study, Furnham and Stringfield (1998) broadened the multi-rater source domain by including consultants as a rater group in a study on congruence in multi-rater assessment. They found that congruence between self-supervisor, self-peer, and self-consultant was very low, whereas congruence between supervisor-peer, supervisor-consultant, and peer-consultant was quite high. Thus, both of these recent studies are consistent with the meta-analysis conducted by Harris and Schaubroeck.

These results support the use of multi-rater assessment in that we are likely to gain a more objective assessment of one's job performance by asking others. As Mabe and West (1982) point out, there is a great deal of reliance on the use of self-evaluations despite the limitations of such an approach. Thus, given these results, self-evaluations alone should be treated cautiously.

Notwithstanding that self-ratings are fraught with biases, it is important to look further for other explanations for lack of agreement between multi-rater sources. Recall earlier that we reviewed the multidimensionality of performance. Perhaps, in addition to the leniency bias of self-raters, the different sources for ratings simply look at different facets of leader performance and therefore report less agreement. Borman (1974) was the first to point out that because raters at

different organizational levels have different interactions with target leaders, they attend to different facets of performance. Finally, Thorton (1980) provides further evidence for this by finding that peer, self, and supervisor ratings did not share similar factor structures.

These results taken as a whole suggest that perhaps it is not conceptually wise to equate each group by having them assess exactly the same dimensions. By doing so, it should come as no huge surprise that there is less than perfect agreement. These sentiments were first offered by Borman (1974) who suggested that raters should be sub-grouped by organizational level with each rater providing performance feedback for a target leader using only dimensions appropriate for their level in the hierarchy. This very same suggestion is again offered by Bozeman (1997) in a commentary on congruence in multi-rater agreement. He suggests that we need to be cautious when considering methodological issues in conjunction with multi-rater assessment given the degree of incongruence between rating groups (peers, self, supervisor, subordinate).

2.5 Stressors

Stressors are organizational or situational conditions that require an individual's adaptive response (e.g., work overload, financial problems, sleep restriction). Thus, stressors can be viewed as stimuli to which an individual must respond. We will now review stressors that we believe will be present in garrison environments. Some stressors examined here will be common to the occupational stressor literature (for an excellent review, see Jex, 1998), whereas some will be specific to the sample and setting. Given this, it is useful to adopt a classification scheme for different types of stressors. For our purposes, we examined three types of stressors that would be subsumed under Jex's (1998) classification scheme: 1) role stressors, 2) interpersonal conflict, and 3) situational constraints.

2.5.1 Role Stressors

A role is a set of behaviors expected of a person in a given position. We all play many roles in our daily lives and work is no exception. Katz and Kahn (1978) point out that in social systems such as organizations roles serve the important function of coordinating individual members' behavior and that we receive role information formally and informally. However, when information needed to fulfill a role is less than clear or a person is mismatched with a given role, roles can be stressful to the individual. A common *role stressor* is role (or work) overload (Jex & Bliese, 1999).

Role overload is a stressor that occurs in the workplace when the work an employee needs to do, or is directed to do, is more than he or she can finish successfully in a given time. Further, the worker may simply perceive the work to be excessively demanding. Jex (1998) points out that it is useful to make the distinction between *quantitative* and *qualitative* overload when conceptualizing role overload. Quantitative overload occurs when an employee has the requisite skills and experience to meet the demands of the job, however, there are too many demands put on the employee at one time to handle. Hence the term quantitative overload; the employee simply cannot attend to all of the role demands due to limited time and resources. In contrast, qualitative overload occurs when the demands of the role exceed an employee's skills and abilities.

In a military unit, either type of role overload may affect performance. A soldier may be faced with what he/she perceives as excessive demands on his/her time to perform the mission/task to standard. For instance, a squad leader may be tasked with multiple assignments and not perceive a way to perform all that has been assigned to standard in the allotted time. Likewise, a soldier may in fact feel that they are "in over their head" when faced with a tasking that either exceeds their current training or ability to perform it to standard. Either conceptualization of role overload might adversely impact a soldier's performance, which in turn could impact the unit since most tasks in the military are additive in nature (e.g., Steiner, 1972). That is, the success of the mission is equal to the sum of individual efforts of all unit members--if a soldier succumbs to role overload then they may not be able to contribute to affect overall unit performance.

2.5.2 Interpersonal Conflict

Because military duty always involve close coordination and mutual support among soldiers and units to complete a given mission, interpersonal conflict can be seen as an especially debilitating stressor of performance in the military. When the pressure to perform as a group is coupled with other stressors in military environments (e.g., danger of work, deployed settings, sleep restriction) even a minor argument among unit members can have a critical impact on an entire unit. Therefore, when performance suffers in the military, it is possible that interpersonal conflict may be a telltale causal agent. For this reason, measuring intra-group (within the unit) conflict becomes important to military researchers as a possible source of stress and its potential impact on the unit's mission.

2.5.3 Situational Constraints

As stated above, the effectiveness of a unit will be a function of the individual efforts of its soldiers. Therefore, it is necessary for a military unit to create as positive an organizational environment as feasible to facilitate individual and unit performance. However, given the unique nature of military duty, many organizational stressors exist and cannot be completely eliminated. In the present study we examined five situational constraints, which can be viewed as stressors in the Army. These are work-family conflict; sleep restriction, work unpredictability, relationship problems, and financial problems.

2.6 Strains (Antecedents of Performance)

Strains can be seen as the response to a stressor, or the resulting negative adaptation stemming from the stressor (e.g., negative well-being, low morale, low job satisfaction, depression, physical symptom logy). Occupational stress models examine strains as outcomes of stressors. There are a wide variety of strains that can be assessed from an organizational theory standpoint (e.g., job satisfaction, turnover, absenteeism). Likewise, there are many strains that can be examined from a health perspective (e.g., distress, depression, physical and psychological well-being). In sum, there is evidence that the health and organizational effects of stressors have important consequences within the organization. The present study studied both organizational and health strains as a function of stressors and in conjunction with performance outcomes.

It is important here to again make note of a key feature of the project model concerning strains. We equate strains to antecedents of performance. Simply put, by equating strains to antecedents

of performance we implicitly suggest that strains are negative adaptive responses that precede performance by "setting the stage" for performance outcomes; the general thesis here being that if strains are high, performance suffers. Exactly how strains may affect performance outcomes warrants some further attention.

One possible way to interpret the link between strains and performance is by making use of some well-known performance models in the industrial and organizational research areas. Researchers have noted that performance is not merely a function of ability. Other factors play a role in determining performance. For example, Muchinsky (1993) suggests that *ability, motivation, and situational factors* interact to either facilitate or inhibit performance (Muchinsky, 1993). Though it is easy to see why ability can be equated to performance, this aspect of performance only tells part of the picture. Even if one's ability is high, this only implies that one is capable of performing well; motivation and situational factors are equally important components in discriminating high/low performing units and/or individuals. Simply put, skill or ability is only a single component of determining performance.

Similarly, Campbell et al.'s (1993) model, noted above, frames this same general notion by identifying *declarative knowledge* (knowledge about facts), *procedural knowledge* (job knowledge/skill) and *motivation* as the key determinants of performance. That is, under Campbell's model ability can be seen as a function of declarative and procedural knowledge. It seems clear that the military relies on its soldiers having a great deal of declarative and procedural knowledge, and this has indeed been addressed by researchers (see Hunter, 1983). This point is well illustrated by viewing the military's emphasis on standard operating procedures (SOPs) in the performance of assigned tasks.

When this emphasis on procedural knowledge, i.e., SOPs, is coupled with the similar training, education, and socialization experiences shared by soldiers, it becomes clear that motivation plays a critical role in our ability to differentiate between characteristics of high/low performing units. Simply stated: if soldiers between units share the same procedural knowledge expertise and similar organizational/social experiences, then motivation within a unit may contribute substantially to its performance.

Therefore, addressing the role that motivation may play in unit performance seems especially important in order to help understand the aspects of performance which can discriminate high performing units from low performing units. The present research will address a number of variables that can be viewed as motivationally based and as important antecedents (determinants) of performance for high performing units. These include: cohesion, well-being, job satisfaction, and morale.

2.7 Measuring Performance

Understanding what performance is and how to measure it may seem straightforward. Unfortunately it is not. Many researchers have lamented that it is the ultimate "criterion problem" with no easy solution. Measuring performance correctly depends on a host of factors including the purpose of its measurement, multidimensionality, and its relatedness to other aspects of performance. Before narrowing our focus on the given sample, we define performance and compare it with other similar constructs.

2.7.1 Performance, Behavior, Effectiveness and Productivity

Campbell (1990) makes distinctions between job performance, job behavior, effectiveness, and productivity. Equating job performance with all behaviors one engages in at work overly simplifies what job performance is. Not all behaviors at work are necessarily related to job performance. That is, not all behaviors at work have an evaluative component, which is necessary when considering performance. Thus, Campbell (1990) suggests that performance "reflects members' contributions to organizational goals--behaviors that lead to or detract from a position's contribution to organizational effectiveness."

Effectiveness also can be seen as a criterion outcome. However, effectiveness includes measures such as salary level, promotion rate and productivity, each of which may be out of the control of the individual or be determined by other things than job performance. Consider a case where an employee is a high performer, yet is not given high marks on effectiveness because of facets out of the control of the individual (e.g., personality conflict with supervisor, lack of proper resources). For this reason, it is useful to differentiate between effectiveness outcomes and performance outcomes.

Lastly, productivity takes into account the benefit/cost relationship. Jex (1998) states that productivity involves cost in that two companies may receive equal benefits (\$) for a specified period of time, but reaching that level of benefit may have cost one company more than the other. In this case, the company that had less cost to get more is seen as more productive. Having differentiated performance from similar constructs we now address how to adequately measure job performance. As noted in our discussion of the 360-degree assessment process, it is widely believed that performance is multidimensional in nature. Therefore, tapping behavioral dimensions that capture all facets of performance becomes critical.

2.7.2 Enlarging the Criterion Space: Dimensions That Matter

Ideally, one would like to have several "objective" measures of performance. That is, measures which unequivocally separate units that are likely to successfully perform their missions from units that are not likely to be successful. There are, however, both practical and conceptual difficulties associated with identifying objective measures of unit performance.

In terms of practicality, it is possible that "objective" performance measures exist in archival forms or can be created using expert observers such as Observer-Controllers (OCs). The two potential problems with this approach are (1) it generally requires considerable manpower to sift through Archival information or to train OCs to be non-biased, reliable raters; and (2) there is reason to believe that archival data may not be reliable. For example, anecdotal information suggests that Quarterly Training Brief (QTB) information does not always reliably differentiate between high and low performing units. Additionally, military researchers have long bemoaned the difficulty of using OCs (e.g., NTC or JRTC) for ratings of performance. This difficulty stems from the OCs primary mission: training and not evaluation (see Hodges, 1994). Its a paradoxical situation since OCs, if allowed to do so, could offer rich data if trained as non-biased and reliable raters. In terms of validation work, unreliable performance data is actually worse

than no performance data. This is because unreliable data can be misleading, while absent data is just absent.

More controversial in terms of "objective" measures is the theoretical notion of what it means for a unit to be effective. For example, gunnery scores might be considered an "objective" performance criterion for an Armor Company, but a measure of this nature is more of a performance measure for a specific task (i.e., gunnery), than a measure of overall unit performance. To be effective, a unit undoubtedly needs to have members who have proficient task mastery (e.g., gunnery skills); however, task mastery alone does not ensure success. Instead, factors such as positive cohesion, effective teamwork, confidence, and physical and mental resilience will weigh heavily on unit success (e.g., Manning & Ingraham, 1987). That is, the social or human dimensions are likely to play a large role in the performance of a unit. Organizational theorists have noted that performance is a function of ability, motivation and a litany of situational factors including individual and group factors (e.g., Muchinsky, 1993). In short, what these issues highlight is that there is really no simple measure of unit performance. To adequately measure performance, one needs to include both hard and soft measures of performance by tapping task proficiency (when available), appropriate archival information (when available), and unit climate information.

The challenge of measuring performance in a valid and comprehensive manner does not belong solely to the military; many organizations are dealing with this issue as well. In fact, the field of performance measurement is currently receiving a great deal of attention in industrial/organizational research (Campbell, 1999). In the current research, a number of behavioral dimensions of performance will be measured based on a model originally proposed by Campbell, McCloy, Oppler & Sager (1993). Thus, it is our intent here to utilize Campbell et al.'s robust model of performance to serve as a framework from which to assess relevant performance dimensions, both "hard" and "soft", in military units--the multidimensional nature of performance. Figure 1.2 below illustrates the performance measures we sought as classified under the Campbell model.

Figure 1.2: Dimensions of Performance Assessed in the current study subsumed under the Campbell Model

<u>Behavioral Dimension of Performance</u>	<u>Unit Archive--"Hard"</u>	<u>Climate Measure--"Soft"</u>
Job-Specific Task Proficiency	ARTEP, Gunnery, Rifle Marksmanship	Combat Readiness Scale
Non-Job Specific Task Proficiency	APFT, SRP	N/A
Demonstration of Effort	Unit Readiness	Commitment Scale
Maintenance of Personal Discipline	UCMJ actions, Police Reports, Sick Call, Profiles, Quarters	OCB Scale, Retention Intent Scale
Facilitation of Peer/Team Performance	Awards, Achievements, Re-enlistment, Badges	OCB Scale
Supervision	Soldier Development	N/A

Note: Campbell et al.'s (1993) also includes Management/Administration and Written/Oral Communication as behavioral dimensions; neither is assessed in the current research. ARTEP = Army Training and Evaluation Program, SRP = Soldier Readiness Program, OCB = Organizational Citizenship Behavior.

2.8 Leader Moderators

Lastly, we felt it was important to address individual leader differences in our model. That is, the degree to which these Company Commanders displayed certain personality traits and how they prioritize certain aspects of performance may impact their 360-degree evaluations, their unit's climate, and their unit's performance.

A key contribution made by industrial/organizational psychology is the advent of the Five-Factor Model (FFM). The FFM, derived using factor analytical techniques across a number of studies and settings, has reliably demonstrated that personality is made up of 5 stable traits: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism. Of these five personality types conscientiousness has emerged as the most valid predictor of job performance (for excellent meta-analytical reviews see Barrick & Mount, 1991 or Salgado, 1997). Thus, we measured the FFM personality trait of conscientiousness along with the FFM personality trait of neuroticism. We felt that since we were examining stressor-strain data in our study that leader neuroticism may play a role in a leaders' ratings and his/her unit's performance.

Additionally, another personality variable often assessed in conjunction with research involving stressor-strain relationships is the Type A behavior pattern (Rosenbaum, 1978). Impatience, hostility, irritability, job involvement, competitiveness, and achievement striving characterize type A individuals. Type A has been studied by researchers interested in understanding its moderating effects on work performance outcomes and support has been found for its role as a moderator (e.g., Lee, Ashford, & Jamieson, 1993; Kushnir & Melamed, 1991). Thus, we included it as a personality measure in the present study.

Lastly, we felt that because these leaders (Company Commanders) were being evaluated using multi-rater assessment and their unit's performance data, it was also necessary to ask them which dimensions of performance they see as important. Manning and Ingraham (1987) first suggested this strategy. In a study examining unit cohesion in military training environments, Manning and Ingraham (1987) suggest that leaders may juggle priorities or even selectively neglect some aspects of performance in lieu of others--depending on each leader's emphasis. If each leader picked priorities slightly different it may account for how their unit performance and how they are rated as leaders.

3 METHOD

3.1 Sample

The sample in the present study consisted of 2,403 US Army soldiers assigned to two different units at two Army posts located in the United States: a Field Artillery Brigade (n = 703) and Armored Cavalry Regiment (n = 1700). In order to examine model variables at the aggregate level, soldiers were nested within 31 company-sized elements--13 companies from a Field Artillery Brigade and 18 companies from an Armored Cavalry Regiment. 91% of the participants were male and 9% were female. Racially, the sample was 50% Caucasian, 18% Hispanic, 16% African-American, 3% Asian, 3% Multi-Racial, and 6% Other. 80% of participating soldiers had high-school diplomas with 15% having an Associate's Degree or some college and 5% having a Bachelor's Degree or higher. 92% of soldier participants were between the ranks of E1 to E6 (Private to Staff Sergeant).

3.2 Procedure

WRAIR and CAL researchers collected data in two phases: 1) with the Field Artillery Brigade (January 2000) and 2) next with the Armored Cavalry Regiment 360-degree leadership assessments (February 2000). Each phase was conducted in a similar fashion as follows.

360-degree leadership assessments were first collected for each company-level commander by the Center for Army Leadership using a rating scheme carefully constructed by CAL researchers and the unit's Personnel staff. This rating scheme made use of the Company Commander's subordinates, peers, supervisors and self-evaluations. Each rating group was given a few days to complete the 360-degree assessments.

There were two minor differences between the 360-degree data collected with the Field Artillery and the Armored Cavalry Regiment. First, paper and pencil surveys were used with the Field Artillery Brigade and web-based surveys were used for the Armored Cavalry Regiment. With the Armored Cavalry Regiment, leadership assessment data were collected via computer on an intranet-based data collection program in the Field Artillery Brigade. Leaders logged on to a central website, identified the leader whom they were rating and their association to that leader (peer, self, superior, subordinate) and completed the feedback survey. Additionally, the response format for each item was changed for the Armored Cavalry Regiment data collection.

Roughly two weeks after the 360-leadership assessments were taken WRAIR researchers collected unit climate data from each of the Company Commanders' units. Climate surveys were

in paper and pencil format using mark-sense forms. These surveys were distributed down to the Company level and administered internally by each unit. Once completed, WRAIR researchers quickly scanned and analyzed the climate data and prepared feedback for the units. Concurrently, WRAIR researchers obtained access to all unit performance archival records and recorded these for use in future analyses. CAL next gave each Company Commander feedback on his/her 360-degree assessments.

During the initial briefing on how the 360-degree assessments would be presented, CAL researchers collected individual difference data from each of Company Commander using a short paper and pencil survey designed by WRAIR. Following this, CAL researchers gave Company Commanders the detailed developmental feedback on the results of their 360-degree assessments in counseling sessions. Immediately following CAL feedback sessions, WRAIR presented unit climate results in a unit-wide briefing. Though these briefing were given to the entire Brigade/Regiment, each Commander was given an anonymous code so results would remain confidential when presented publicly.

Following the briefings given by CAL and WRAIR researchers, WRAIR personnel collected and organized all unit performance data. Performance data were subsequently matched with all other study data (360-degree assessments, individual differences, stressors & strains) and were then aggregated to the company level for analysis.

3.3 2.3 Scales and Data used to Assess Model Variables

3.3.1 360-degree Leadership Ratings--CALs Azimuth

The Azimuth is a multi-rater assessment tool developed by the Center for Army Leadership. Because the Azimuth is still under development, some items were changed from the initial data collection with the Field Artillery Brigade and the final data collection with the Armored Cavalry Regiment. Thus, in the present study we used items used in both data collections. Similarly, the response format was changed from an 11-point format (1 = Never; 11 = Always) in the Field Artillery data collection to a 5-point response format (1 = Never, 5 = Always) in the Armored Cavalry Regiment data collection. Therefore, it was necessary to standardize (z-score) these items within each sample before combining both samples for analyses.

The Azimuth measures 16 dimensions of leadership based on the Army's Be-Know-Do Model of leadership (see FM 22-100, Military Leadership). Therefore, these behavioral dimensions of leadership were based on endorsed Army doctrine. Refer to Appendix A for a summary of these dimensions. Appendix B provides the items used to tap each dimension in the present study. We report means and standard deviations for each Azimuth Dimension per rating group (see Table 1) (self, peer, subordinate, supervisor), though for brevity's sake, we will only use global ratings (e.g., a single self rating value made up of the sum of all self dimension rating values) in our analyses. Lastly, we will report agreement/congruence between the rating groups.

Recall that these multi-rater leadership assessments are at the individual level of analysis while stressors, strains, and performance were measured at the company level of analysis. Hence we tie a leader to their unit's climate and performance.

3.3.2 Individual Difference Moderators

We noted above that we also asked each leader, i.e., the 31 Company Commanders to respond to a short survey measuring individual differences that may moderate stressor-strain and strain-performance relations. Individual difference measures included the personality measures of Type A (Spence, Helmreich & Pred, 1987), conscientiousness and neuroticism (Goldberg, 1992) and the rank ordering of each dimension of performance as organized under the Campbell Model.

3.3.3 Stressors

Stressors assessed were both scale-based and one-item self-report in nature. Stressors measured with scales were workload (perceived role overload, Cammann et al., 1983), Work-Family Conflict (Netemeyer, Boles & McMurrian, 1996), and Intra-Group Conflict (modified from Spector & Jex, 1999). One-item self-report items assessed in the study measured sleep restriction, work, predictability, relationship problems, and financial problems. WRAIR Researchers have used these items in past stress research. Aggregate level means, standard deviations and alphas are reported in Table 2.

3.3.4 Strains

Strains/Performance Antecedents assessed were all scale based in nature. Strains/Performance Antecedents measured were Psychological Well-Being (Goldberg, 1972), Job Satisfaction (developed at WRAIR, similar to the Job Diagnostic Survey General Satisfaction Scale, Hackman & Oldham, 1975), Unit Cohesion (WRAIR Scales), Vertical Cohesion for Officers & Non-Commissioned Officers (both WRAIR Scales), and Morale (WRAIR Scale). Note that Unit Cohesion and Vertical Cohesion are frequently treated as moderators in the WRAIR Stressor-Strain-Performance model (e.g., Bliese & Castro, 2000). In this report, however, we lack sufficient power to adequately test moderator models; consequently, we examine these variables as strains or outcomes. Aggregate-level means, standard deviations, and alphas are reported in Table 2. These data were aggregated to the unit level.

3.3.5 Archival Performance and Scale-Based Performance Measures

In order to assess the multidimensional nature of performance we gathered both archival performance data from each company (N = 31) involved in the study. Additionally, we assessed aspects of performance via scale-based methodology. In order to structure our efforts to collect objective unit performance measures using archival records, we made use of the Campbell model previously discussed. However, we were only successful in collecting archival performance data on some of the dimensions proposed under the original framework. Means and standard deviations for the archival and survey-based performance measures we were able to collect are summarized in Table 2. Again, these are reported at the aggregate level.

3.4 Plan for Data Analyses

Because this study was exploratory in nature, we sought only to test the relationships between 360-degree assessments, our climate data, and archival performance using correlational analyses.

In particular, we tested zero-order correlations at the group-level (31 companies) for each unit and its commander. Thus, statistics reported are aggregated. In addition to correlational analyses, an additional objective was to examine the psychometric properties of CALs leadership assessment tool, Azimuth. Therefore, we examined the psychometric properties of the Azimuth by looking at the factor structure of the scale and by examining congruence in ratings between the multiple sources.

4 RESULTS

4.1 Means and Standard Deviations for all Study Variables

Means and standard deviations for the aggregate measures employed in the study are summarized in Tables 1-3: Table 1 contains descriptive statistics for the stressors studied; Table 2 contains descriptive statistics for the strains/performance antecedents; and Table 3 contains descriptive statistics for the performance measures.

Table 4 presents means and standard deviations for all Azimuth dimensions of leadership assessed by the rating sources. Note that these scores were standardized and in the case of subordinate and peer ratings were aggregated.

Unfortunately, leader personality traits of interest in the study could not be examined as only 12 of 31 company commanders responded to the individual difference survey. Because of the poor return rate, we were not able to address the role of leader differences in personality and prioritizing of performance dimensions in conjunction with any of our analyses because the sample size was too small for any generalizations to be made.

4.2 Correlational Analyses: Stressors, Strains, and Performance

4.2.1 Stressors and Strains

Our first objective was to test the link between stressors, strains and performance (both archival performance measures and survey-based performance measures). We used Pearson correlation coefficients to test these relationships. Since our study was based at the company level, the relationships were tested at the company level resulting in a sample size of 31.

Table 5 presents the results of the relationships between stressors and strains. Notice first that there were fairly strong relationships among stressors. That is, units that had high levels on any one specific stressor also tended to have high levels on other stressors. Thus:

- *Unit stressors are highly correlated with each other.*

Second, notice that most strains (morale, job satisfaction, well-being) were highly correlated with each other. The exception to this rule was horizontal cohesion. Horizontal cohesion was unrelated to any of the other strains. This is somewhat of an atypical finding as horizontal cohesion is generally positively related to vertical cohesion. Nonetheless one can conclude that:

- *Unit strains are highly correlated with each other.*

The aspect of Table 5 that is of most interest is the relationships between stressors and strains. Notice that, overall, stressors were significantly correlated with strains such that units with high stressors had low well-being, morale, job satisfaction, etc. The only exception to this rule was that none of the stressors were related to horizontal cohesion. Once again, we find this to be an atypical finding, but may reflect the fact that horizontal cohesion is more often conceptualized as a moderator variable than as an outcome. In general, however, we can conclude that:

- *Units with high stressors tend to have poor vertical cohesion, poor well-being, poor morale, and poor job satisfaction.*

The stressor that was most strongly related to strains (as a whole) was intra-group conflict (based on an average correlation). In other words,

- *The amount of intra-group conflict that a unit reported was most predictive of the levels of strain in the unit.*

4.2.2 Stressors and Performance

Table 6 presents the results for the relationships between stressors and unit performance measures. Four of the unit performance measures are survey based. These are:

- Affective Commitment
- Continuance Commitment
- Organizational Citizenship Behaviors
- Perceptions of Combat Readiness

The remaining three performance measures are archival based.

- M16 Scores
- APFT Scores
- UCMJ Offenses

Correlations among performance measures. As noted previously, the stressors tend to be correlated, such that a unit that is high on one stressor will, in general, be high on the other stressors. In contrast, notice that the relationship among the performance outcomes tends to be less clear. For instance, units that have high M16 scores also tend to have high APFT scores ($r=.68$); low UCMJ violations ($r=-.43$); and high citizenship behavior ($r=.57$). This is an expected series of relationships. Interestingly, however, units with high M16 scores also tended to have low continuance commitment ($r=-.36$). This is not an expected relationship. Similarly, units that report high combat readiness tend to report high affective commitment ($r=.38$) and high citizenship behavior ($r=.44$). Interestingly, though, readiness is not significantly related to M16 scores, APFT scores or UCMJ violations. One would expect units that have high readiness scores to have high M16 and APFT scores.

Finally notice that units that have high APFT scores tend to report high citizenship behaviors ($r=.40$); however, they also report low continuance commitment ($r=-.66$) which is not expected.

From a modeling perspective, the results do suggest that organizational citizenship behaviors (OCBs) are a useful survey-based performance dimension that are related to archival-based performance measures. Specifically, units high on OCBs tended to have high M16 Scores, high APFT Scores, and low (but non-significant) UCMJ violations. Thus we conclude that:

- *OCBs are an important performance dimension that can be assessed via survey instrument and that are related to archival performance measures.*

Relationships among stressors and performance. One of the key objectives of this research was to examine the relationship among stressors and performance indices. These relationships are also presented in the box within Table 6.

Eight of the 42 correlations in the box are significant. Based on a p-value of .05, we would expect about two correlations to be significant merely on a chance basis. So, overall, there is evidence to suggest that stressors are related to performance. Interestingly, six of the eight significant correlations involved either affective commitment or continuance commitment. High work overload, high work-family conflict, high intra-group conflict and high marital problems were correlated with low affective and continuance commitment. Thus we can conclude that:

- *Work stressors have a large impact on soldiers' commitment to the Army, such that high stress results in low continuance and affective commitment.*

There were two significant correlations that did not involve commitment. The first correlation was between work-family conflict and APFT scores. The results revealed that:

- *Units with high work-family conflict tend to have high APFT scores.*

It is unclear why this relationship occurs. It is possible that the relationship exists because units with high APFT scores are spending more time on PT, and this time takes away from family.

The second significant correlation was between financial problems and UCMJ violations such that,

- *Units where members had a lot of financial problem stress tended to have a high number of UCMJ violations.*

Based on the results as a whole, we conclude that the main effect of high stress environments is to reduce soldier's commitment to the Army. Interestingly, however, units with low commitment may not necessarily have performance in other domain areas. Recall that units with high APFT scores tended to have low continuance commitment ($r=-0.66$). Similarly, units with high M16 scores tended to have low continuance commitment ($r=-0.36$). These results highlight the fact that performance is a multidimensional domain, and increases in one aspect of performance may be accompanied by decreases in other areas of performance.

4.2.3 Strains and Performance

The box in Table 7 provides the results for the relationship between strains and performance. Eleven of the 42 correlations were significant.

Five of the significant correlations in Table 7 involved affective commitment. All of the strains except horizontal cohesion were significantly related to affective commitment. The highest correlation was between NCO vertical cohesion and affective commitment. This suggests that soldiers felt a strong attachment to the Army when they perceived that their NCOs were considerate and competent. Note also that NCO vertical cohesion was related to continuance commitment while officer vertical cohesion was not. This again suggests that NCO consideration is particularly important in terms of soldier commitment.

The second performance dimension that was highly related to the strains was combat readiness. Unit cohesion, NCO vertical cohesion and officer vertical cohesion were all significantly related to reports of readiness. Officer vertical cohesion was the variable most strongly related to readiness with a correlation of 0.76.

From these correlations we can conclude that:

- *Perceptions that officers are considerate and competent is particularly important in terms of perceptions of readiness; while perceptions that NCOs are considerate and competent is particularly important in terms of commitment.*

The final two significant correlations involved organizational citizenship behaviors (OCBs). Officer vertical cohesion and overall morale were both related to OCBs. It would be logical to conclude that:

- *Considerate and competent officer leadership in a unit is associated with high morale and high citizenship behaviors.*

Finally, notice that none of the strains were significantly related to archival performance measures such as M16 scores, APFT scores or UCMJ violations.

4.3 Exploratory Factor Analysis: The Factor Structure of the Azimuth

Our second objective was to test the psychometric properties of CALs 360 assessment tool, the Azimuth. We did this by conducting an exploratory factor analysis to determine the factor structure of the Azimuth. Recall that CAL designed the Azimuth to tap into the themes from the Be-Know-Do Model of Leadership. Furthermore, WRAIR researchers were interested in uncovering other aspects of leadership not tapped by their current measures. That is, the Azimuth may provide insight into other factors of leadership that need to be assessed. Results from factor analysis revealed that the items were too highly related to differentiate any distinct factors. In fact, only one "meta" factor emerged from a scale with 62 items. This "meta" factor accounted for 94% of the variance in Azimuth items and had an overall eigenvalue of 50.66.

Limitations and suggestions given the results of the factor analysis will be taken up in the discussion.

4.4 Correlational Analyses: Rater Congruence, Ratings and Performance

Our final objective was to test the link between 360-degree assessments and performance (unit objective measures and survey-based measures). In order to test these links, we calculated Pearson correlation coefficients. As in the previous analysis, these relationships were analyzed at the aggregate level. However, we opted not to examine each 360-degree dimension assessed in the Azimuth. We instead made use of global ratings for each rating source. For example we computed Pearson correlation coefficients between subordinate ratings and combat readiness, or Army Physical Fitness Scores. These relationships are summarized in Table 8.

Rating source intercorrelations. In Table 8, notice the intercorrelations among leadership ratings. Two ratings sources were significantly related. Specifically, self-ratings were significantly positively related to supervisor ratings ($r=0.52$), and Supervisor ratings were significantly related to subordinate ratings ($r=0.74$). None of the other correlations were significant. This suggests that supervisors and subordinates tend to be very similar in their perceptions of the company commanders' leadership. Company commanders also tend to have ratings similar to their supervisors' ratings.

Leadership ratings and performance. The leadership ratings that was most related to the performance outcomes were the ratings from peers. Peer ratings were positively related to (a) M16 scores, (b) APFT scores and (c) organizational citizenship behavior ratings. The other leadership ratings were significantly related to only one performance outcome -- either combat readiness or citizenship behaviors. Specifically, self-ratings were related to OCBs; and supervisory and subordinate ratings were related to reported readiness.

These findings are interesting, for they suggest that commanders tended to feel that they were good leaders if their company members reported high citizenship behavior. In contrast, supervisors and subordinates tended to report that the company commander was a good leader if the unit members (as a whole) reported high combat readiness. Finally, peers reported that a company commander was a good leader if the company commander's unit had good M16 scores, high APFT scores, and high citizenship behavior. This suggests that different targets were using different criteria to evaluate leadership.

5 DISCUSSION

5.1 Summary of findings

The present study had three objectives:

- To assess the impact of unit climate (unit stressors and strains) on both survey-based and archival-based measures of unit performance.

- To determine the relationship between multi-rater leadership assessments (360 degree assessments) and unit performance.
- To validate the Center for Army Leadership's 360-degree Azimuth tool by examining the psychometric properties of CALs Azimuth.

5.1.1 Unit Climate and Unit Performance

In this research we used a very broad definition of unit performance. This definition included archival-based performance measures (PT scores, UCMJ violations, M16 Scores) as well as survey-based measures (commitment, citizenship behaviors, and perceptions of unit readiness). As stated above, we were interested in determining the relationship between unit climate factors (stressors and strains) and unit performance measures.

Climate and Archival Performance. In general, we found that unit climate measures were poor predictors of archival performance measures. That is, we found only two significant correlations out of 36 possible. These two significant correlations were between Work-Family conflict and APFT Scores ($r = 0.41$) and between Financial Problems and UCMJ Offenses ($r=0.34$). It seems logical that units that had high APFT scores might also have high Work-Family conflict since high APFT scores may be a sign that soldiers are spending a lot of time on PT. Similarly it seems logical that units with members who, on average, have financial problems would be more likely to have UCMJ violations. Presumably, financial problems drive soldiers to engage in behaviors that lead to UCMJ violations.

Despite the fact that the two relationships appear plausible, we recommend exercising caution in interpreting them. This is because one would expect, by chance, to find 2 significant correlations in a matrix of 36 correlations even if there were no meaningful relationships. Thus, it is possible that the two significant correlations are "spurious". That is, they reflect random chance more than any true relationship. The only way to verify these results would be to cross-validate them on another sample.

Does this mean that unit climate has no relationship with archival performance measures? While it appears that there is no relationship, we believe that it is more likely that a relationship exists, but that we lacked the statistical power in this sample to detect relationships. Recall that we had only 31 units even though we had over 2500 individual respondents. It is common for relationships among data from different sources (archival versus survey) to be relatively small (see Thomas, Dickson, & Bliese, in press). Thus, it will be necessary to continue to build up the sample sizes to truly determine the links between unit climate measures and archival performance measures.

Climate and Survey-Based Performance Measures. In contrast to the situation involving archival measures, we found quite a large number of significant relationships among unit climate measures and survey-based performance measures. In total there were 48 correlations involving stressors/strains and survey-based performance measures. Of these 48 correlations, 17 were significant -- many more than would be expected by chance. Interestingly, 12 of the 17 correlations involved one or the other form of commitment (affective or continuance). Units that

had high stressors (work overload, work family-conflict, high intra-group conflict) tended to have members with low commitment to the Army. Similarly, units with high strains (poor well-being, poor morale, low satisfaction) also tended to have members with low commitment. In the WRAIR stressor-strain-performance model we interpret these findings as evidence that high work stress leads to high strains, and high strains, in turn, lead to low commitment to the Army. It is interesting to note that Unit Cohesion and Vertical Cohesion were also related to commitment. In the WRAIR stressor-strain-performance model, cohesion is conceptualized as a moderator (see Bliese & Jex, 1999). Nonetheless, it is clear that these "moderators" have a direct impact. Note, in particular, that the item with the strongest relationship to affective commitment was NCO vertical cohesion. This implies that the biggest predictor of soldiers' commitment to the Army is their perception of the consideration and competence of their NCOs.

Two other survey-based performance measures merit discussion. Three unit-level factors were related to combat readiness: (a) Unit Cohesion, (b) Vertical Cohesion from NCOs, and (c) Vertical Cohesion from officers. Of these three, officer vertical cohesion had the strongest relationship with reported combat readiness ($r=0.76$). This finding in conjunction with the commitment findings suggest that officer leadership is instrumental in terms of unit combat readiness, but NCO leadership is instrumental in terms of soldiers attraction to and commitment to the Army.

Finally, the results involving Organizational Citizenship Behaviors (OCBs) were encouraging. OCBs are behaviors that are not part of any official job title, but are behaviors that individuals engage in that nonetheless help the organization perform well. These are behaviors such as "Volunteer to do things without being asked", "Help others who have heavy workloads", "Assist unit leaders with their work." Presumably, units high in OCBs have the potential to be high performers. The two significant predictors of high OCBs in a unit were (a) Vertical Cohesion from officers, and (b) overall Morale. Once again, it appears that officer leadership is instrumental in OCBs -- a key performance dimension.

5.1.2 360 leadership ratings and unit performance

The second key goal of this research was to examine relationships among multi-rater leadership assessments (360 degree assessments) and unit performance. There were 28 possible correlations involving leadership assessments and unit performance. Of these, six were significant. By chance one would expect one or two significant correlations.

The rating source that had the most significant correlations with unit performance was peer ratings. Peer ratings were related to (a) M16 Scores, (b) APFT Scores, and (c) OCBs. This is an interesting finding, because the causal nature of these relationships is difficult to discern. On the one hand, it is possible that leaders who are identified by their peers as being "good" also have units with high M16 scores: high APFT scores, and high OCBs.

It is also possible, however, that Quarterly Training Briefs (QTBs) and other formal evaluations make M16 and APFT scores particularly salient. These are scores that are easily quantifiable

and understood, which makes them an appropriate and natural gauge by which leaders rate their peers. As such, peers may simply conclude that leaders with high M16 and APFT scores are "good". In any event, it is clear that peers focus on unit OCBs, M16 scores and APFT scores when rating each other. Note that no other rating source appeared to use archival information when evaluating leaders.

Each of the three remaining rating sources was significantly correlated with one performance dimension. Specifically, self-ratings were significantly correlated with OCBs; and Supervisory and Subordinate ratings were significantly correlated with Combat Readiness.

In summary, supervisors and subordinates tended to rate the company commander highly when the commander's unit reported high combat readiness. In contrast, individual company commanders tended to rate themselves highly if their unit had high OCBs. Finally, peers' ratings were related to unit OCB, unit M16 scores and unit APFT scores.

In terms of congruence, the results showed that supervisor ratings tended to be correlated with both self and subordinate ratings. Peer ratings were not related to any other ratings.

5.1.3 Psychometric properties of the Azimuth

The final goal was to validate the Center for Army Leadership's 360-degree Azimuth tool by examining the psychometric properties of CALs Azimuth. The analyses revealed that there was little discriminant validity among the scale in the Azimuth. Instead, raters tended to use a global response format -- if they believed that the company commander was an effective leader, they rated him or her highly on all dimensions. In contrast, if they believed that the company commander was a poor leader, they rated him or her low on all dimensions.

Clearly the lack of dimensionality could limit our ability to detect relationships. One would expect that specific leadership behaviors would be related to specific performance dimensions. Thus, in the absence of psychometrically distinct specific leadership attributes it may be difficult to detect relationships. Future research should consider refining the Azimuth. One possibility would be to target specific questions to specific rating groups, since different groups see and respond to different leadership dimensions.

5.2 Future Work

While this work was non conclusive in many respects, it nonetheless adds to our understanding of the relationships among stressors-strains and performance. We can conclude that positive garrison environments characterized by relatively low work stress and effective leadership play a big role in soldier commitment. NCO leadership, in particular, is important to soldier commitment.

The data also showed that leadership is related to performance dimensions. In the analyses we merged data from multiple sources and showed that ratings of unit effectiveness (reports of combat readiness and organizational citizenship behaviors) were related to independent leadership ratings of the company commanders.

The work of science is typically one of continual refinement and improvement. Future studies should build on this work by including a larger number of groups to enhance the statistical power. In addition, the Azimuth should be modified to allow for better measures of specific dimensions of leadership. Changes such as these will allow for a more thorough examination of the relationships between (a) unit climate and unit performance, and (b) leadership and unit performance.

Note that this data can also serve as an excellent data source for modeling individual outcomes. The analyses reported in this report focused on the company-level results, and this resulted in a lack of power. Analyses focused on individual-level models, in contrast, have considerable power. Bliese, Ritzer, Thomas and Jex (2000), for example, conducted an in-depth analysis of how individual and group-level factors impact commitment.

6 REFERENCES

- Barrick, M. R., & Mount, M. K. (1991). The "Big Five" personality dimension and job performance: A meta-analysis. Personnel Psychology, 44, 1-26.
- Bernardin, H. J. & Beatty, R. W. (1984). Performance Appraisal: Assessing Human Behavior at Work. Boston, MA: Kent Publishing Company.
- Bettenhausen, K.L. & Fedor, D. B. (1997). Peer and upward appraisals: A comparison of their benefits and problems. Group & Organization Management, 22, 236-263.
- Bliese, P. D. & Castro, C. A. (2000). Role clarity, work overload and organizational support: Multilevel evidence of the importance of support. Work and Stress, 14, 65-73.
- Bliese, P. D., & Jex S. M. (1999). Incorporating multiple levels of analysis into occupational stress research. Work and Stress, 13, 1-6.
- Bliese, P. D., Ritzer, D., Thomas, J. L. & Jex, S. M. (2000). Supervisory Support and Stressor-Commitment Relations: A Multi-Level Investigation. Manuscript in Review.
- Borman, W. C. (1991). In M. D. Dunnette & L. M. Hough (Eds.), Handbook of Industrial and Organizational Psychology, 2nd. (pp. 271-326). Palo Alto, CA: Consulting Psychologists Press, Inc.
- Bozeman, D. P. (1997). Inter-rater agreement in multi-source performance appraisal: a commentary. Journal of Organizational Behavior, 18, 313-316.
- Cammann, C., Fichman, M., Jenkins, G. D., & Klesh, J. (1983). Michigan organizational assessment questionnaire. In S. E. Seashore, E. E. Lawler, P. H. Mirvis, and C. Camman (Eds.), Assessing Organizational Change: A Guide to Methods, Measures, and Practices, New York: Wiley-Interscience, 71-138.
- Campbell, J. P. (1990). Modeling the performance prediction problem in industrial and organizational psychology. In M. D. Dunnette & L. M. Hough (Eds.), Handbook of industrial and organizational psychology (2nd ed., Vol. 1, pp. 687-732). Palo Alto, CA: Consulting Psychologists Press.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993) A theory of performance. In N. Schmitt & W. C. Borman (Eds.), Personnel selection in organizations (pp.35-70). San Francisco: Jossey-Bass.
- Campbell, J. P. (1999). The definition and measurement of performance in the new age. In D. R. Ilgen & E. D. Pulakos (Eds.), The changing nature of performance: Implications for staffing, motivation, and development (pp. 39-429). San Francisco: Jossey-Bass.
- Cascio, W. F., & Bernardin, H. J. (1981). Implications of performance appraisal litigation for personnel decisions. Personnel Psychology, 34, 211-226.
- Fiske, S. T., & Taylor, S. E. (1991). Social cognition (2nd edition). New York: McGraw-Hill.
- Furnham, A., & Stringfield, P. (1998). Congruence in Job-Performance Ratings: A study of 360-degree feedback examining self, manager, peers, and consultant ratings. Human Relations, 51, 517-530.
- Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness. London: Maudsley Monographs.

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. J. Deary, F. De Fruyt, & F. Ostendorf (Eds.) Personality Psychology in Europe, Vol. 7, pp. 7-28. Tilburg University Press: The Netherlands.

Hackman, J. R. & Oldham, G. R. (1975). Development of the Job Diagnostic Survey. Journal of Applied Psychology, 60, 159-170.

Harris, M. H., & Schaubroeck, J. (1988). A meta-Analysis of self-supervisor, self-peer, and peer-supervisor ratings, Personnel Psychology, 41, 43-62

Hodges, J. S. (1994). Analytical use of data from Army training exercises: A case study in tactical reconnaissance. Journal of the American Statistical Association, 89, 444-451,

Holzbach, R. L. (1978). Rater bias in performance ratings: Superior, self, and peer ratings. Journal of Applied Psychology, 63, 579-588.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors across studies. Psychological Bulletin, 96, 72-98.

Jex, S. M. (1998). Stress and job performance: Theory, research, and implications for managerial practice. Thousand Oaks, CA: Sage

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. W. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, NJ: General Learning Press

Katz, D., & Kahn, R. L. (1978). The social psychology of organizations (2nd ed.). New York: John Wiley

Kushnir, T., & Melamed, S. (1991). Work-load, perceived control and psychological distress in Type A/B industrial workers. Journal of Organizational Behavior, 12, 155-168.

Latham, G. P. & Wexley, K. N. (1982). Increasing productivity through performance appraisal. Reading, Mass.: Addison-Wesley.

Lee, C., Ashford, S. J. & Jamieson, L. F. (1993). The effects of Type A behavior dimensions and optimism on coping strategy, health and performance. Journal of Organizational Behavior, 14, 143-157.

London, M., & Smither, J. W. (1995). Can multi-source feedback change perceptions of goal accomplishment, self-evaluations, and performance-related outcomes? Theory-based applications and directions for research. Personnel Psychology, 48, 803-839.

Mabe, P. A. III., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. Journal of Applied Psychology, 67, 280-296.

Manning, F. J., & Ingraham, L. H. (1987). An investigation into the value of cohesion in peacetime. In G. Belenky (Ed.), Contemporary studies in combat psychiatry (pp. 453-470). Westport, CT: Greenwood Press.

Mount, M. K., Judge, T. A., Scullen, S. A., Systma, M. R., Hezlett, S. A. (1998). Trait, rater, and level effects in 360-degree performance ratings. Personnel Psychology, 51, 557-576.

Muchinsky, P. M. (1993). Psychology applied to work. (4th ed.). Pacific Grove, CA: Brooks/Cole.

Netemeyer, R. G., Boles, J. S. & McMurrian, R. (1996). Development and validation of work-family conflict and family-work conflict scales. Journal of Applied Psychology, 81, 400-410.

Rosenman, R. H. (1978). The interview method of assessment of the coronary-prone behavior pattern. In Dembroski, T. M., et al. (eds.), Coronary Prone Behavior, (pp. 55-69). Spring-Verlag: New York.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 10, pp. 173-220). San Diego, CA: Academic Press.

Salgado, J. F. (1997). The five-factor model of personality and job performance in the European community. Journal of Applied Psychology, 82, 30-42.

Spector, P.E., & Jex, S. M. (1998). Development of Four Self Report Measures of Job Stressors and Strain: Interpersonal Conflict at Work Scale, Organizational Constraints Scale, Quantitative Workload Inventory, and Physical Symptoms Inventory. Journal of Occupational Health Psychology, 3, 356-367.

Spence, J. T., Helmreich, R. L., & Pred, R. S. (1987). Impatience versus achievement strivings in the Type A pattern: Differential effects on student's health and academic achievement. Journal of Applied Psychology, 72, 522-528.

Steel, R. P., & Ovalle, N. K. (1984). Self-appraisal based upon supervisory feedback. Personnel Psychology, 37, 667-685.

Steiner, I. D. (1972). Group process and productivity: Orlando, FL: Academic Press.

Thomas, J. L. (1999). Personality and motivational predictors of military leadership assessments in the U.S. Army Reserve Officer Training Corps. Unpublished doctoral dissertation, Wayne State University, Detroit, Michigan.

Thomas, J. L., Dickson, M. W., & Bliese, P. D. (in press) Values Predicting Leader Performance in the US Army Reserve Officer Training Corps Assessment Center: Evidence for a Personality-Mediated Model. Leadership Quarterly.

Thorton, G. C. (1980). Psychometric properties of self-appraisals of job performance. Personnel Psychology, 33, 414-432.

US Army (1990). FM 22-100: Military Leadership.

US Army (1997). Officer Evaluation Reporting System: Army Regulation 623-105

Zammuto, R. F., London, M., & Rowland, K. M. (1982). Organization and rater differences in performance appraisals. Personnel Psychology, 35, 643-657.

7 TABLES

7.1 Table 1: Descriptive Statistics for Aggregate Level Study Stressors (n = 31 Companies).

Stressor	M	S.D.
Workload	3.09	.20
Work-Family Conflict	3.02	.29
Intra-group Conflict	2.93	.20
Sleep Restriction	2.80	.20
Financial Problems	3.01	.23
Relationship Problems	2.98	.27

Note. Workload, Work-Family Conflict, & Intra-group Conflict are scale-based stressors. In order to place them in a common metric, scores for each were summed then divided by the number of items. Sleep Restriction, Financial Problems, and Relationship Problems are one-item measures.

7.2 Table 2: Descriptive Statistics for Aggregate Level Strains/Performance Antecedents (n = 31 Companies).

Strain	M	S.D.
Unit Cohesion	3.40	.14
Vertical Cohesion	3.24	.27
Officer		
Vertical Cohesion	3.41	.23
NCO		
Well-Being	2.99	.09
Job Satisfaction	3.13	.22
Morale	3.24	.20

Note. All strain/performance antecedent measures are scale-based. In order to place them in a common metric, scores for each were summed then divided by the number of items.

7.3 Table 3: Descriptive Statistics for Aggregate Level Performance Dimensions (n = 31).

Strain	M	S.D.
M16 Qualifying Scores	32.59	3.14
Army Physical Fitness Test Scores	252.82	9.65
UCMJ Offenses in Unit	12.50	9.29
Continuance Commitment	2.54	.17
Affective Commitment	2.94	.18
Org. Citizenship Behavior	2.43	.14
Combat Readiness	3.28	.29

Note. M16 Scores, APFT Scores, and UCMJ Offenses are archival performance dimensions and the remaining strain/performance antecedent measures are scale-based. In order to place them in a common metric, scale-based scores were summed then divided by the number of items.

7.4 Table 4: Means and Standard Deviations (Standardized) for all 360-degree Azimuth Dimension x Rating Source

360-Degree Dimension	Self Rating	Supervisor Rating	Peer Rating	Subordinate Rating
Communicating	-.03 .68	-.09 .71	-.06 .86	-.04 .91
Decision- Making	.10 .50	-.18 .73	.09 .81	-.11 .88
Motivating	-.01 .57	-.03 .70	-.05 1.00	-.03 .86
Developing	.08 .59	.14 .65	-.10 1.05	-.10 .84
Building	.07 .68	-.12 .85	-.15 .82	-.03 .86
Learning	.06 .67	.03 .71	-.23 .82	-.06 .89
Planning & Organizing	-.08 .72	.09 .68	-.11 .89	-.08 .89
Executing	.02 .77	-.20 .85	-.04 .81	-.04 .87
Assessing	-.05 .93	-.01 .62	.07 .84	.04 .87
Respect	-.08 .81	-.25 .80	-.19 .89	.05 .83
Selfless Service	.13 .82	.07 .74	-.17 .85	-.03 .81
Integrity	.29 .58	-.07 .68	-.15 .89	-.15 .95
Physical Fitness	-.21 .80	.00 .60	-.27 .90	-.11 .96
Tactical	-.05 .79	-.16 .76	-.11 .75	-.04 .82
Technical	-.04 .61	-.16 .55	.04 .71	.01 .91
Overall Leadership	.23 .51	-.20 .84	-.11 .68	.11 .91

Note. Standardized means are in bold print, standard deviations are in normal print.

7.5 Table 5: Correlations between Aggregate Level Stressors (1-6) and Strains (7-12) (n = 31).

Variable	1	2	3	4	5	6	7	8	9	10	11	12
1 Work Overload	--											
2 W-F Conflict	*.85	--										
3 I-G Conflict	*.53	*.51	--									
4 Sleep Restriction	*.53	*.64	.32	--								
5 Financial Problems	*.46	*.52	.23	*.68	--							
6 Relationship Problems	*.40	*.61	.20	*.52	*.59	--						
7 Cohesion	.07	.25	.15	.15	.29	.24	--					
8 VC-NCO	*.38	-.28	*.68	-.23	-.11	-.05	.00	--				
9 VC-OFF	*.44	-.25	*.36	-.11	-.14	.07	.15	*.60	--			
10 Well Being	*.59	*.72	*.53	*.71	*.56	*.53	-.05	*.49	*.43	--		
11 Morale	*.52	*.53	*.46	*.44	-.29	-.22	.04	*.62	*.72	*.74	--	
12 Job Satisfaction	*.53	*.58	*.72	*.55	-.33	-.21	.01	*.56	*.39	*.76	*.65	--

Note. W-F Conflict = Work Family Conflict, I-G Conflict = Intra-Group Conflict, VC-NCO = Vertical Cohesion-Non Commissioned Officers, VC-OFF = Vertical Cohesion-Officers.

* **p < .01--in bold text**; *p < .05 in normal text

7.6 Table 6: Correlations between Aggregate Level Stressors (1-6) and Performance (7-13) (n = 31)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Work Overload	--												
2 W-F Conflict	*.85	--											
3 I-G Conflict	*.53	*.51	--										
4 Sleep Restriction	*.53	*.64	.32	--									
5 Financial Problems	*.46	*.52	.23	*.68	--								
6 Marital Problems	*.40	*.62	.20	*.52	*.59	--							
7 M16 Scores	-.04	.13	.15	-.16	-.26	-.01	--						
8 APFT Scores	.14	*.41	.20	.08	-.06	.18	*.68	--					
9 UCMJ Offenses	.15	.09	.13	.21	*.34	.14	*.43	-.19	--				
10 Affect. Commit.	*.39	*.40	*.56	-.19	-.18	-.29	-.26	-.18	-.15	--			
11 Cont. Commit.	*.41	*.50	-.32	-.08	-.02	*.34	*.36	*.66	-.14	*.52	--		
12 OCB	-.27	-.01	-.22	.00	-.07	.17	*.57	*.40	-.23	.24	-.04	--	
13 Combat Readiness	-.22	-.05	-.14	.03	.00	.11	-.24	-.04	.11	*.38	.21	*.44	--

Note. W-F Conflict = Work Family Conflict, I-G Conflict = Intra-Group Conflict, APFT Scores = Army Physical Fitness Test Scores, Affect. Commit = Affective Commitment, Cont. Commit = Continuance Commitment, OCB = Organizational Citizenship Behavior
*** p < .01--in bold text; *p < .05 in normal text**

7.7 Table 7: Correlations between Aggregate Level Strains and Performance (n = 31)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13
1 Unit Cohesion	--												
2 VC NCO	.00	--											
3 VC Officer	.15	*.61	--										
4 Well-Being	-.05	*.49	*.43	--									
5 Overall Morale	.04	*.62	*.72	*.74	--								
6 Job Satisfaction	.01	*.56	*.39	*.76	*.65	--							
7 M16 Scores	-.27	-.14	-.21	-.15	.03	-.26	--						
8 APFT Scores	.14	-.13	-.02	-.21	.02	-.28	*.68	--					
9 UCMJ Offenses	.31	-.13	-.11	-.10	-.10	-.00	*.43	-.19	--				
10 Affect. Commit.	.10	*.75	*.47	*.57	*.66	*.57	-.26	-.18	-.15	--			
11 Cont. Commit.	-.14	*.40	.15	.23	.13	.20	*.36	*.66	-.14	*.52	--		
12 OCB	.11	.33	*.43	.15	*.38	.12	*.57	*.40	-.23	.24	-.04	--	
13 Combat Readiness	*.44	*.43	*.76	.26	.31	.31	-.24	-.04	.11	*.38	.21	*.44	--

Note. VC-NCO = Vertical Cohesion-Non-Commissioned Officer, VC-Officer = Vertical Cohesion-Officer, Job Satis. = Job Satisfaction, APFT Scores = Army Physical Fitness Test Scores, Affect. Commit = Affective Commitment, Cont. Commit = Continuance Commitment, OCB = Organizational Citizenship Behavior

* **p < .01--in bold text**; *p < .05 in normal text

7.8 Table 8: Correlations between Global 360-degree Assessment by Rating Source and Performance (n=31)

Variable	1	2	3	4	5	6	7	8	9	10	11
1 Self Ratings	--										
2 Supervisor Ratings	*.52	--									
3 Peer Ratings	-.02	.31	--								
4 Subord. Ratings	.26	*.74	.11	--							
5 M16 Score	.33	.06	*.63	-.12	--						
6 APFT Score	.14	.05	*.48	.04	*.68	--					
7 UCMJ Offenses	-.23	.24	.08	.06	*-.43	-.19	--				
8 Affect. Commit	-.22	-.31	-.21	-.25	-.26	-.18	-.15	--			
9 Contin. Commit.	.05	.01	-.27	-.11	-.37	*-.66	-.14	*.52	--		
10 OCB	*.44	.30	*.52	.17	*.57	*.40	-.23	.24	-.04	--	
11 Combat Readiness	.29	*.39	.04	*.50	-.24	-.04	.11	*.38	.21	*.44	--

Note. APFT Score = Army Physical Fitness Test Scores, Affect. Commit. = Affective Commitment, Contin. Commit. = Continuance Commitment, OCB = Organizational Citizenship Behavior.

*** p < .01--in bold text; *p < .05 in normal text**

***Leadership Dimensions Assessed Under the
Be-Know-Do Model (FM 22-100)***

<i>Leaders of Character and Competence</i>			<i>Acts to achieve excellence by providing purpose, direction, and motivation</i>		
VALUES "Be"	ATTRIBUTES "Be"	SKILLS "Know"	ACTIONS "Do"		
			Influencing	Operating	Improving
Respect	Physical	Technical	Communicating	Planning	Developing
Selfless Service		Tactical	Decision- Making	Executing	Building
Integrity			Motivating	Assessing	Learning

9 APPENDIX B: Items from CAL Azimuth 360-Degree Assessment Tool

Communicating

Provides clear direction.
Explains own ideas so they are easily understood.
Keeps others well-informed.
Listens well.
Tells it like it is.

Decision-Making

Delays decisions unnecessarily (r).
Generates innovative solutions to unique problems.
Makes sound decisions.
Involves others in decisions that affect them.

Motivating

Creates a supportive environment.
Disciplines in a firm, fair, and consistent manner.
Inspires people to do their best.
Often acknowledges good performance of others.
Sets clear performance expectations.

Developing

Encourages professional growth.
Is an effective teacher.
Uses counseling to provide performance feedback.
Provides opportunities to learn.
Seldom delegates authority (r).

Building

Actively participates in unit activities.
Encourages cooperation among team members.
Focuses the unit on mission accomplishment.
Treats others as valuable team members.

Learning

Becomes defensive when given critical feedback (r).
Encourages open discussion to improve unit.
Helps the unit adapt to changing circumstances.
Seems to be realistic about own personal limitations.
Willingly accepts new challenges.

Planning & Organizing

Reasonably adheres to 1/3 2/3 rule in mission planning.
Develops effective plans to achieve the units' goals.
Sufficiently allocates appropriate resources.
Sets clear priorities.
Unwilling to modify original plan when circumstances change (r).

Executing

Completes assigned mission to standards.
Does not meet mission timelines (r).
Does what ever is necessary (within ethical limits) to complete the mission.
Monitors execution of plans to identify problems.

Assessing

Rarely conducts after action reviews (r).
Takes time to find out what others are doing.

Respect

Is discrete/tactful when correcting or questions others.
Creates a climate of fairness and equal opportunity in the unit.
Checks on the safety and well-being of others.
Treats others with respect.

Selfless Service

Accepts the blame for the team.
Places welfare of others above personal gain.
Shares hardships with others.

Integrity

Demonstrates moral courage (does what is right).
Is honest in word and deed.

Physical

Maintains a high level of energy under adverse conditions.
Demonstrates an appropriate level of fitness.

Tactical

Possesses the necessary technical expertise to accomplish assigned tasks.
Refines plans to exploit unforeseen opportunities.

Technical

Applies the right tactics, techniques, and procedures in consideration of METT-TC.

Overall Leadership

This person is someone I would follow into combat.